

## **Impact of Multi-Source Feedback on Leadership Competency Development: A Longitudinal Field Study\***

Guangrong Dai

*Research Scientist*

Korn/Ferry Leadership and Talent Consulting

Kenneth P. De Meuse

*Associate Vice President of Research*

Korn/Ferry Leadership and Talent Consulting

Clarke Peterson

*Principal Consultant*

Atlanta Leadership Consulting, LLC

Multi-source feedback interventions have become widespread in many organizations around the world (Brutus *et al.*, 2006). It has been reported that multi-source feedback is used by 90% of the *Fortune* 1000 companies in the United States and 85% of Australia's top 500 corporations (Carruthers, 2003). The popularity of such a human resource practice has stimulated much research enthusiasm in the academic field. According to Morgeson, Mumford, and Campion (2005), more than 100 multi-source feedback related articles have been published since 1990. In general, the practice of providing multi-source feedback is viewed favorably, with researchers espousing the potential benefits for performance improvement at both the individual and organizational level (Bailey and Fletcher, 2002; Edwards and Ewen, 1996; Fletcher and Baldry, 1999; London and Smither, 1995). However, empirical research demonstrating the effectiveness of multi-source feedback is sketchy and mixed. For example, Kluger and DeNisi (1996) asserted that the link between providing feedback and performance improvement is not uniform. In fact, it has been shown that multi-source feedback can have a negative effect on performance (see Pfau and Kay, 2002). A recent meta-analysis reported an

---

\*A version of this paper was presented at the 2007 Annual Meeting of the Academy of Management in Philadelphia.

overall positive but very small effect size of multi-source feedback effectiveness (Smither *et al.*, 2005). Based on this research, these authors have argued that practitioners should not expect large, widespread performance improvement after employees receive multi-source feedback. This article focuses on a specific type of multi-source feedback – a competency-based feedback program – and examine whether it helps managers develop their leadership competencies through a longitudinal study.

## MULTI-SOURCE FEEDBACK EFFECTIVENESS

### Evaluation of the Effects of Multi-Source Feedback

Researchers have investigated the impact of multi-source feedback for decades. For instance, Seifert, Yukl, and McDonald (2003) identified 14 field studies dating back to the 1970s (e.g., Hegarty, 1974). The criterion in these studies usually was a change in ratings of managerial behavior or skill. However, a review of the extant literature revealed that virtually all research studies investigated the performance improvement by comparing the *average* or *composite* performance or skill rating at Time 1 (before feedback) to the one at Time 2 (after feedback; e.g., see Johnson and Ferstl, 1999; Reilly *et al.*, 1996; Seifert *et al.*, 2003; Smither *et al.*, 2003). Smither *et al.* (2005) suggested that this comparison may be too crude of a measure of behavioral change. Many organizations use managerial performance multi-source feedback for leadership development. Such developmental programs generally recommend that feedback recipients should focus on no more than two or three leadership areas for improvement (Antonioni, 1996). Indeed, in many feedback programs, managers select a small number of improvement goals. For example, many organizations provide managers multi-source feedback on a whole host of leadership competencies. As such, managers usually select to work on only a few bottom-ranked (but important) competencies for development. These managers might make substantial improvement on these selected leadership competencies, but this gain likely would have little effect on *overall* leadership as indicated by a composite score. Hence, the amount of improvement observed in the literature may not accurately reflect the true success of developmental multi-source feedback. For this reason, the primary purpose of the current study is to investigate the improvement on leadership competencies recipients specifically selected for development following multi-source feedback. It is predicted that feedback recipients will make significantly greater gains on the selected leadership competencies than on the other competencies or on their overall leadership score.

### Longitudinal Effects of Multi-Source Feedback

The cost of implementing a multi-source feedback program tends to be quite expensive. It is a critical issue for organizations to know whether such a program leads to improvement and whether the improvement is sustained over a long period of time. If feedback recipients return to their previous level of competency, the substantial financial investment made in such programs will be of little value (Reilly *et al.*, 1996). Walker and Smither (1999) followed 252 managers over multiple administrations of an *upward feedback* program. They found no improvement between the first and second feedback, but they did observe improvement between the second and the third and again between the third and fourth administration of the upward feedback. Consequently, time of

assessment may be an important consideration to document the multi-source feedback effectiveness. As previously mentioned, feedback recipients typically select to develop their bottom-ranked competencies in managerial performance feedback programs. One might contend that the improvement observed after the initial administration of the feedback program is due to regression to the mean. Thus, a second purpose of the current study is to examine whether the initial gains following the first administration of the multi-source feedback were sustained (i.e., across multiple administrations of feedback). If the improvement persists and is greater than the amount that can be attributed to regression to the mean, then there is more confidence in declaring the effectiveness of the multi-source feedback program.

In the context of providing feedback, a truly longitudinal study will be the one that includes more than one assessment after the initial feedback. Unfortunately, there have been only a few longitudinal studies published in the literature (e.g., see Walker and Smither, 1999; Reilly *et al.*, 1996), and all of them have obtained ratings from only a single source (i.e., direct reports). Individuals' reactions to feedback likely are sensitive to rater dynamics (Bailey and Austin, 2006). For example, how different rater groups (such as boss, peers, and direct reports) agree with each other on their ratings may influence the feedback recipients' initial reactions to the feedback, which in turn could affect subsequent motivational attitudes and behaviors (Smither *et al.*, 2005). Hence, the current study investigates the developmental outcomes within the context of 360-degree feedback solicited from several rater groups.

### **Influence of the Developmental Difficulty of Leadership Competencies**

Smither *et al.* (2005) proposed a theoretical model for multi-source feedback and performance improvement. In this model, they theorized that eight factors play a role in behavioral change and performance improvement following multi-source feedback. These eight factors fall into three different categories: (1) characteristics of the feedback itself, such as self-other agreement (Atwater and Yammarino, 1992; Fletcher and Baldry, 2000; Johnson and Ferstl, 1999) as well as the favorability of the feedback (Atwater and Brett, 2005); (2) characteristics of the feedback recipient, such as self-efficacy (Bailey and Austin, 2006) and goal orientation (Heslin and Latham, 2004); and (3) characteristics of the rater source, such as the credibility of the raters (Albright and Levy, 1995). The present study postulates that another factor may be missing from the model – the nature of the performance dimension or leadership competency managers select to improve.

The nature of the performance dimension itself appears likely to directly or indirectly influence feedback effectiveness. Not all performance dimensions (i.e., competencies) are equal in cognitive complexity, experience requirements, and emotional involvement. Some areas of performance are harder to improve than others. The nature of the performance dimension may directly impact the feedback effectiveness, in that feedback recipients likely will make more improvement on easier performance dimensions than harder ones. Further, the nature of the performance dimension may interact with other factors to influence feedback effectiveness indirectly. For example, Bailey and Austin (2006) examined how perceived importance of the feedback behaviors to the individual's current role affects performance improvement based on the logic that the perceived importance is related to feedback salience. Another example of the indirect effects of the nature of the performance dimension on feedback effectiveness is its relationship with self-other agreement. Self-other agreement was found to be influenced by performance

dimension rating difficulty (Wohlers and London, 1989). The credibility of rater source also may interact with the performance dimension. For example, a rating on the competency "developing direct reports" is likely to be perceived more credible from direct reports than from other rater sources (e.g., the boss and peers).

Consequently, the final objective of the current study is to investigate the potential impact of the inherent nature of the performance dimension on the results of multi-source feedback. In this study, the variable to be investigated is the developmental difficulty of the leadership competencies. Developmental difficulty appears to be an obvious factor that influences feedback effectiveness. While logically sound, it has not been empirically examined. It is predicted that feedback recipients will make more improvement on relatively easier-to-develop leadership competencies than on harder ones.

Before proceeding to the description of the research method, the practice of "competency modeling" is briefly reviewed here. Competency modeling has been widely embraced by practitioners during the past few years. The Job Analysis and Competency Modeling Task Force sponsored by the Society for Industrial and Organizational Psychology conducted an extensive literature review and interviewed 37 subject matter experts from varying backgrounds asking how they would define a competency (Schippmann *et al.*, 2000). There was a wide range of definitions. Some of the experts perceived that competencies were not fundamentally different from traditionally defined KSAOs (knowledge, skills, abilities, and other characteristics). In contrast, some authors viewed competencies as a combination of several individual characteristics (e.g., motives, traits, values, knowledge, and skills) which can be shown to differentiate superior from average performers (Spencer *et al.*, 1994).

In the leadership field, a few implicit assumptions seem to be fundamental to competency modeling. First, managerial performance can be organized into a number of performance taxonomies. These performance taxonomies provide a basis for detailed comparisons among managerial jobs and the people in them (Tett *et al.*, 2000). Second, successful performance on each of the managerial performance taxonomies can be attributed to a cluster of attributes, characteristics, or qualities (Lombardo and Eichinger, 2004; Spencer *et al.*, 1994). These clusters of people capabilities can be referred to as "leadership competencies." Thus, a competency is a cluster of personal attributes which are required for successful performance on a managerial or leadership performance dimension. It is believed that there is a great deal of similarity among competencies between management functions, across organizations, and throughout various levels of management (Schippmann *et al.*, 2000). Competency modeling provides a common language to describe leadership capabilities and the job requirements for management positions (Lievens *et al.*, 2004). In addition, competency modeling ties the derivation of management job profiles to the organization's strategy. The emphasis on strategic human resource management has made competency based practices popular around the world (Becker *et al.*, 2001).

## METHOD

### Participants and Organizational Context

Seventy-eight managers from a large financial services company located in the United States participated in this study. The company initiated a leadership development

program in 1997 in an effort to improve the quality of its current management team and to enhance bench strength. Although this program is ongoing, the last data collection for the current study was 2003. The central purpose of the program was to provide 360-degree feedback to managers, implementing individual development plans (IDPs) which focus on competencies in need of development for either the current job role or for some future role.

All participating managers had at least one "development cycle." A development cycle began with one multi-source feedback (pre-assessment) and ended with another multi-source feedback (post-assessment). Typically, the time interval between assessments was one and two years. The variation occurred, because the managers' had different business agendas. For each administration of the multi-source feedback, raters provided both *skill* and *importance* ratings. Between the two feedback sessions, participants worked with an internal executive coach. All managers had the same executive coach; he was the Senior Vice President of Human Resources and Organizational Development at this company. The coaching included such activities as facilitating the feedback, discussing things with which managers were struggling, identifying developmental needs, selecting competencies for development, creating action plans, and meeting from time to time to review progress. The selected competencies usually were from the bottom-ranked on *skill* ratings but top-ranked on *importance* ratings. The IDPs generally consisted of 10-15 pages and specified several potential resources that would help the manager's development. Additionally, each manager was provided a book containing many practical suggestions on how to develop specific leadership competencies (Lombardo and Eichinger, 1996).

The company held managers accountable to complete the work outlined in their development plans in a number of ways. For instance, managers were asked to supply a progress report regularly. All reports were read and graded. Substantive responses were written and provided to each manager. Most managers received a grade of "B," meaning that they were approaching their plan well and making good progress. Those managers earning a "C" were given some suggestions, and it was made clear that they were not making sufficient progress. To receive an "A," managers had to do something out of the ordinary. These individuals would receive a congratulatory e-mail with a copy going to their bosses up to and including the CEO. Perhaps, most importantly, 20% of the manager's bonus was at stake – dependent on the satisfactory completion of the plan within the specified time period. The collection of the progress reports, grading them, and then directly tying a significant portion of compensation to their completion were designed to clearly demonstrate the importance of leadership development.

Fifty-five (71%) of the 78 managers had experienced two development cycles. As a result, these managers experienced three administrations of the multi-source feedback. Sixteen (21%) had three development cycles. These 16 managers experienced four administrations of the multi-source feedback. For managers who had gone through more than one development cycle, nearly all of them had selected and worked on different leadership competencies during each development cycle. In total, less than 7% of the time the competencies selected were the same over multiple development cycles. For each administration of the multi-source feedback, each of the managers had been rated by eight people (i.e., their immediate boss, three peers, three direct reports, and self). In order to maintain respondent confidentiality, no information on age, ethnicity, or gender was collected.

### Multi-Source Feedback Instrument

The multi-source feedback instrument used in the study consisted of 67 items designed to assess six factors and 67 leadership competencies. The theoretical foundation and research on the development of this instrument are summarized in Lombardo and Eichinger (2003, 2004). Several other researchers have used the Lombardo and Eichinger competency model and assessment in their studies (e.g., Lievens *et al.*, 2004; Tett *et al.*, 2000).

Each of the leadership competencies was rated on a 5-point scale to assess how *skilled* the manager was on that competency. Responses ranged from “a towering strength” (5), “talented” (4), “skilled/ok” (3), “a weakness” (2), to “a serious issue” (1). A skilled and unskilled behavioral description was provided for each of the competencies to facilitate the rating. According to Lombardo and Eichinger (2003), test-retest reliability was 0.75 for self-ratings and 0.82 for others’ ratings (i.e., boss, peers, and direct reports). Internal consistency (Cronbach alpha) ranged from 0.77 to 0.93 for the six factors. In addition, each of the leadership competencies was rated on a 5-point scale to assess how *important* it was for the manager’s job. The scale ranged from “mission critical” (5), “very important” (4), “nice to have” (3), “less important” (2), to “least important” (1).

### Developmental Difficulty

The instrument publisher has developed an index of “developmental difficulty” to indicate how hard it is to develop each of the 67 leadership competencies (for details, see Lombardo and Eichinger, 1995). This index is based upon the instrument developers’ evaluation of the *underlying nature of the competency*. The underlying nature of the competency consists of six factors – complexity of the skills involved, experience requirements, beliefs (e.g., attitudes, values, and opinions), cognitive complexity, emotion involvement, and human make-up. Lombardo and Eichinger coded each of the competencies along the six factors using a 5-point scale, ranging from “easiest” (1) to “most difficult” (5). Thus, each competency had six difficulty scores. An overall difficulty index then was computed for each competency by adding the above components. Subsequently, the 67 competencies were grouped into the following five categories based upon the distribution of the overall difficulty index: (a) “easiest” (1); (b) “easier” (2); (c) “moderate” (3); (d) “harder” (4); and (e) “hardest” (5). The publisher of this instrument provided no empirical data related to the developmental index (Lombardo and Eichinger, 1995). However, both authors have extensive experience and are widely cited for their research in the field of leadership development (see Lombardo and Eichinger, 1989, 2004). Table 1 depicts each of the 67 competencies and their corresponding developmental difficulty index score and category.

### Analysis

The pre-post averages of the “all others” ratings (i.e., the manager’s boss, peers, and direct reports) were compared to measure the improvement in each development cycle. Self-ratings were not included in the analysis, because research suggests that they are less accurate relative to the ratings from others (Dunning *et al.*, 2004; Eichinger and Lombardo, 2004). The scores from the different rater groups were consolidated to enhance the reliability of the assessment. Ratings from the “peers” and the “direct

**Table 1**  
**Developmental Difficulty for Leadership Competencies**

<b>Leadership Competency</b>	<b>Developmental Difficulty Index Score</b>	<b>Difficulty Category</b>
1. Action Oriented	17	1
2. <i>Dealing with</i> Ambiguity	28	4
3. Approachability	24	3
4. Boss Relationships	25	3
5. Business Acumen	23	3
6. Career Ambition	24	3
7. Caring About Direct Reports	24	3
8. Comfort Around Higher Management	26	3
9. Command Skills	24	3
10. Compassion	28	4
11. Composure	27	4
12. Conflict Management	32	5
13. Confronting Direct Reports	27	4
14. Creativity	25	3
15. Customer Focus	18	1
16. <i>Timely</i> Decision Making	18	1
17. Decision Quality	20	2
18. Delegation	21	2
19. Developing Direct Reports & Others	27	4
20. Directing Others	21	2
21. <i>Managing</i> Diversity	28	4
22. Ethics & Values	26	3
23. Fairness to Direct Reports	22	2
24. Functional/Technical Skills	13	1
25. Hiring & Staffing	24	3
26. Humor	24	3
27. Informing	17	1
28. Innovation Management	29	5
29. Integrity & Trust	20	2
30. Intellectual Horsepower	22	2
31. Interpersonal Savvy	28	4
32. Learning on the Fly	24	3
33. Listening	20	2
34. Managerial Courage	29	5

Leadership Competency	Developmental Difficulty Index Score	Difficulty Category
35. Managing & Measuring Work	20	2
36. Motivating Others	26	3
37. Negotiating	27	4
38. Organizational Agility	28	4
39. Organizing	20	3
40. <i>Dealing with</i> Paradox	28	4
41. Patience	23	3
42. Peer Relationships	22	2
43. Perseverance	18	1
44. Personal Disclosure	27	4
45. Personal Learning	30	5
46. Perspective	25	3
47. Planning	18	1
48. Political Savvy	32	5
49. Presentation Skills	24	3
50. Priority Setting	20	2
51. Problem Solving	24	3
52. Process Management	24	3
53. <i>Drive for</i> Results	19	2
54. Self-Development	24	3
55. Self-Knowledge	25	3
56. Sizing Up People	27	4
57. Standing Alone	23	3
58. Strategic Agility	27	4
59. <i>Managing Through</i> Systems	28	4
60. <i>Building Effective</i> Teams	28	4
61. Technical Learning	20	2
62. Time Management	19	2
63. Total Work Systems	19	2
64. Understanding Others	34	5
65. <i>Managing</i> Vision and Purpose	25	3
66. Work/Life Balance	28	4
67. Written Communications	21	2

*Note.* This list of competencies is copyrighted by Lominger International: A Korn/Ferry Company. © 1995. All rights reserved. Do not duplicate, copy, or create derivatives, in any form, without express written permission. For the development difficulty category, 1 = Easiest, 2 = Easier, 3 = Moderate, 4 = Harder, 5 = Hardest.



reports" rater sources were first averaged. The score on a competency represents the mean of the three rater source averages (peers, direct reports, and boss). Subsequently, the improvement score (i.e., the difference between post- and pre-scores) was correlated with the coding of the development difficulty to test whether the recipients' improvement was influenced by the developmental difficulty.

It is acknowledged that multi-source feedback has relatively low inter-rater agreement (Harris and Schaubroeck, 1988; Lawler, 1967). Traditionally, this discrepancy has been attributed largely to rater perspectives that are associated with the different rater groups, such as boss, peers, and direct reports (Borman, 1997). However, research indicates that rating variations resulting from rater groups are not necessarily larger than those differences related to individual raters within the same group (Mount *et al.*, 1998; Scullen *et al.*, 2000). Consequently, the calculation of an average score across raters and across rater groups provides a comprehensive assessment of a manager's leadership performance and significantly reduces rating bias and random error (Scullen *et al.*, 2000).

Across the multiple development cycles, approximately 7% of the time the same competencies were selected by the participants for development. Only the initial improvement was used to determine how it was related to the development difficulty. In total, 54 of the 67 competencies were selected for development. Six of them were selected once. The other 48 were selected by more than one participant. When a competency was selected by more than one participant, the average improvement was used. This approach was employed for two different reasons. First, the average score was a more reliable estimation of the amount of improvement people were likely to make on a competency. Second, from a statistic perspective, the scores were combined into a single number to remove the competency sampling redundancy.

Since there was a variation on the length of a development cycle for different managers, it was necessary to determine that the conclusions made from this study would not be confounded by the differences in the time interval. It was found that time between feedback sessions was not related to the observed pre-post differences on the competencies planned for development ( $r = -0.04$ ,  $p = 0.74$ ), the competencies *not* planned for development ( $r = 0.02$ ,  $p = 0.86$ ), or for the average of *all* competencies ( $r = 0.01$ ,  $p = 0.92$ ). Therefore, the variation on the time interval did not impact the results. It is not to say that the length of time spent on development is unimportant in multi-source feedback. Managers in the current study had a minimum of one year to develop their competencies. As readers will discover later, much of the improvement tends to occur early on in the developmental cycle. Consequently, the addition of time for leadership development made very little difference.

An interesting question was whether participants had a preference toward choosing the easier-to-develop competencies. Of the five difficulty levels of competencies, the percentages of selection (from easiest to hardest) were 11.3%, 26.3%, 29.0%, 24.7%, and 8.7%. The distribution of the development difficulty of the selected competencies and total 67 competencies was not statistically different ( $\chi^2 = 0.8413$ ,  $df = 4$ ,  $p = 0.93$ ), suggesting that participants either were not aware of the developmental difficulties of the competencies, or they selected the competencies without considering the development difficulty.

Finally, it should be noted that the sample size decreased across the three development cycles. One might infer that participants who participated in *all* three development cycles were more motivated toward the leadership development program than other participants. In fact, the sample size decreased mostly due to the way the leadership

development program had been implemented. Managers had different business agendas. The multi-source feedback did not begin at the same time for all the participants. Some managers started early; others participated in the program later. At the time the data were collected for this study, some managers completed three development cycles, some two, and some only had completed one. Consequently, the decrease of the sample size was not due to the attrition of participants. Nevertheless, before the data were analyzed to address the research questions, the impact of the decrease in sample size on the effectiveness of the feedback program was examined.

Among the 78 participants, 16 went through all three development cycles, 39 did two cycles, and the remaining 21 experienced only the first development cycle. After the initial feedback, all participants selected some competencies for development. The performance changes on selected competencies (i.e., differences between the initial rating and the rating after the feedback) were compared for the three groups of participants. Two-way repeated measure analysis indicated that the change on the scores was not statistically different for the three groups ( $F(2, 75) = 0.17, p = 0.85$ ). Similarly, after the second feedback session, the change on the selected competencies was not statistically different for the remaining two groups ( $F(1, 53) = 0.32, p = 0.57$ ). Therefore, the “attrition” of the sample size did not influence the observed improvement on competency ratings. As a result, the three groups of participants were combined for analyses to examine the relevant research questions.

RESULTS

The results of the study are presented below in three sections pertaining to the three development cycles.

First Development Cycle

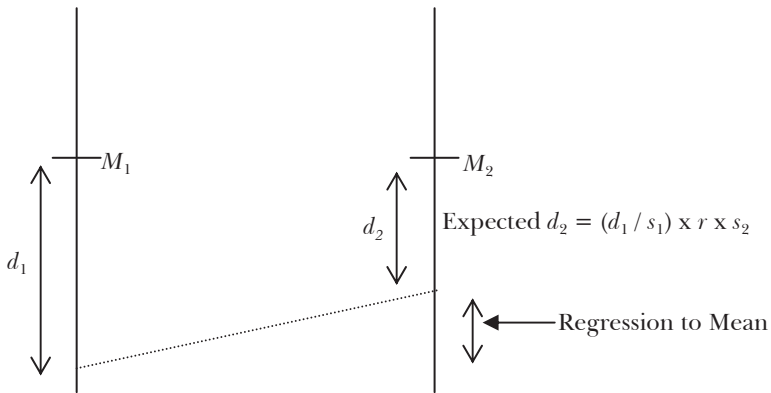
For each manager, the following leadership competency scores were computed: (1) the average score on the competencies selected for development; (2) the average score on the competencies *not* selected for development; and (3) an overall mean composite score across all 67 competencies. The immediate pre- and post-assessment means and standard deviations of the three types of scores in the first development cycle are presented in Table 2.

Table 2  
Immediate Pre- and Post-assessment in First Development Cycle

	Pre		Post		<i>d</i>	<i>p</i> -value
	<i>Mean</i>	<i>Std.</i>	<i>Mean</i>	<i>Std.</i>		
Competencies Selected	3.25	0.35	3.59	0.36	0.94	$p < 0.001$
Competencies Not Selected	3.71	0.22	3.80	0.25	0.38	$p < 0.001$
All 67 Competencies	3.69	0.22	3.79	0.25	0.42	$p < 0.001$

*Note.*  $N = 78$ . *d* denotes effect size, the standardized mean difference.

**Figure I**  
**Regression to Mean**



As can be seen, all three types of averages were statistically different between the pre- and post-administration. The results indicated that managers significantly improved, even on those competencies that they did *not* select for development. However, it should be noted that the managers improved much more on the competencies selected ( $d = 0.94$ ) than on competencies *not* selected ( $d = 0.38$ ), as well as on the overall competency composite score ( $d = 0.42$ ).

As expected, the average score of the competencies selected for development ( $M = 3.25$ ) was statistically lower than the average score of the competencies *not* selected for development ( $M = 3.71$ ,  $p < 0.05$ ). Consequently, it was necessary to examine whether the improvement on the competencies selected for development was due to the regression to mean. Regression to the mean occurs when the scores on those bottom-ranked competencies increase from Time 1 to Time 2 because of a non-intervention related factor; namely, the unreliability of the assessment. In other words, the differences between the bottom-ranked competencies and overall mean on the post-assessment are smaller than the differences between the same competencies and overall mean on the pre-assessment, even when participants did not do anything in particular to develop these competencies. See Figure I for an illustration.

In order to estimate the amount of increase that would occur solely due to regression to the mean, an estimate of the reliability of the multi-source ratings was needed. The correlation between ratings at Time 1 and Time 2 (i.e., coefficient of stability) was used as the estimate of the reliability. This estimate is conservative because the correlation between Time 1 and Time 2 ratings would be less than perfect to the extent that some managers actually improved their performance on some competencies. The correlation between the pre- and the post-assessment scores for each of the participants was computed. The expected difference between the mean of the competencies selected for development and the overall mean on the post-assessment then was computed for each of the participants using the following formula:

$$\text{Expected } d \text{ at Time 2} = (\text{Observed } d \text{ at Time 1} / s_1) \times r \times s_2$$

where observed  $d$  at Time 1 is the difference between overall mean and the mean of the competencies selected for development at Time 1;  $r$  is the correlation between the pre- and post-assessment;  $s_1$  and  $s_2$  are standard deviations at Time 1 and Time 2.

A paired-sample t-test indicated that the observed difference at Time 2 ( $M = 0.19, s = 0.22$ ) was statistically smaller than the expected difference at Time 1 ( $M = 0.25, s = 0.15, p < 0.05$ ). This finding suggests that the observed improvement on the competencies selected for development was significantly greater than explained by regression to the mean.

As mentioned, 55 of the 78 managers had three administrations of the multi-source feedback; 16 of them had four administrations. Table 3 shows the means and standard deviations of the ratings on the competencies selected for development across multiple administrations of the multi-source feedback for the two groups of managers. A one-way ANOVA indicated a significant main effect of time for the group with three administrations ( $F(2, 108) = 66.17, p < 0.001$ ) and the group with four administrations ( $F(3, 45) = 18.31, p < 0.001$ ). The results of *post hoc* paired t-tests revealed that for participants with three administrations, the Post 1 mean rating on the selected competencies ( $M = 3.52$ ) was significantly higher than the Pre mean rating ( $M = 3.18, p < 0.001$ ), and the Post 2 mean rating ( $M = 3.69$ ) was significantly higher than the Post 1 mean rating ( $M = 3.52, p < 0.001$ ). For participants with four administrations, the Post 1 mean rating on the selected competencies likewise was significantly higher than Pre mean rating ( $p < 0.001$ ), and the Post 2 mean rating was significantly higher than the Post 1 mean rating ( $p < 0.001$ ). Although the Post 3 mean rating was higher than the Post 2 mean rating, it was not statistically significant ( $M = 3.75$  and  $3.72$ , respectively;  $p = 0.10$ ). Overall, the improvement in leadership that occurred in the first development cycle continued in the Post 1 and 2 periods, but this improvement leveled off in the Post 3 time frame. Figure II depicts the mean ratings on the competencies selected for development in the first development cycle across multiple administrations of the multi-source feedback.

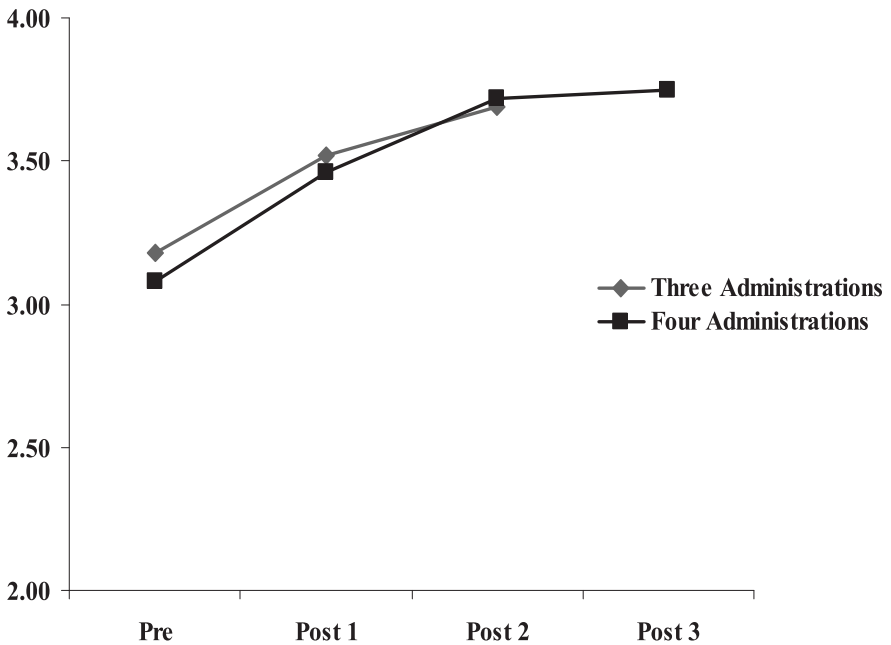
Second Development Cycle

The means and standard deviations for the immediate pre- and post-assessment in the second development cycle are presented in Table 4. All three types of averages are statistically different between the pre- and the post-assessment. The managers improved substantially more on competencies selected for development ( $d = 0.97$ ) than on competencies *not* selected ( $d = 0.38$ ) or on the overall composite score ( $d = 0.38$ ). Again, tests were conducted on whether the improvement could be explained by the

Table 3  
Ratings on Competencies Selected in First Development Cycle

	Pre	Post 1	Post 2	Post 3
Three Administrations ( <i>N</i> = 55)	3.18 ( <i>s</i> = 0.34)	3.52 ( <i>s</i> = 0.34)	3.69 ( <i>s</i> = 0.38)	
Four Administrations ( <i>N</i> = 16)	3.08 ( <i>s</i> = 0.35)	3.46 ( <i>s</i> = 0.31)	3.72 ( <i>s</i> = 0.38)	3.75 ( <i>s</i> = 0.29)

**Figure II**  
**Mean Ratings on Competencies Selected in First Development Cycle**



regression to the mean. The correlations between the pre- and post-assessment scores for the 55 participating managers were calculated. The expected difference between the mean rating of the competencies selected for development and the overall mean rating on the 67 competencies was calculated using the formula described previously. The paired-sample t-test indicated that the observed difference ( $M = 0.20, s = 0.20$ ) was smaller than the expected difference ( $M = 0.24, s = 0.12, p < 0.05$ ), suggesting that the improvement occurring during the second development cycle was greater than simply regression to the mean.

**Table 4**  
**Immediate Pre- and Post-assessment in Second Development Cycle**

	Pre		Post		<i>d</i>	<i>p</i> -value
	<i>Mean</i>	<i>Std.</i>	<i>Mean</i>	<i>Std.</i>		
Competencies Selected	3.31	0.30	3.63	0.34	0.97	$p < 0.001$
Competencies Not Selected	3.75	0.23	3.84	0.25	0.38	$p < 0.01$
All 67 Competencies	3.74	0.23	3.83	0.25	0.38	$p < 0.01$

*Note.*  $N = 55$ . *d* denotes effect size, the standardized mean difference.

**Table 5**  
**Ratings on Competencies Selected in Second Development Cycle**  
**Across Multiple Administrations of the Feedback**

	Pre 1	Pre 2	Post 1	Post 2
Three Administrations (N = 55)	3.37 ( <i>s</i> = 0.29)	3.31 ( <i>s</i> = 0.30)	3.63 ( <i>s</i> = 0.34)	
Four Administrations (N = 16)	3.40 ( <i>s</i> = 0.27)	3.27 ( <i>s</i> = 0.27)	3.70 ( <i>s</i> = 0.17)	3.73 ( <i>s</i> = 0.22)

**Figure III**  
**Mean Ratings on Competencies Selected in Second Development Cycle**

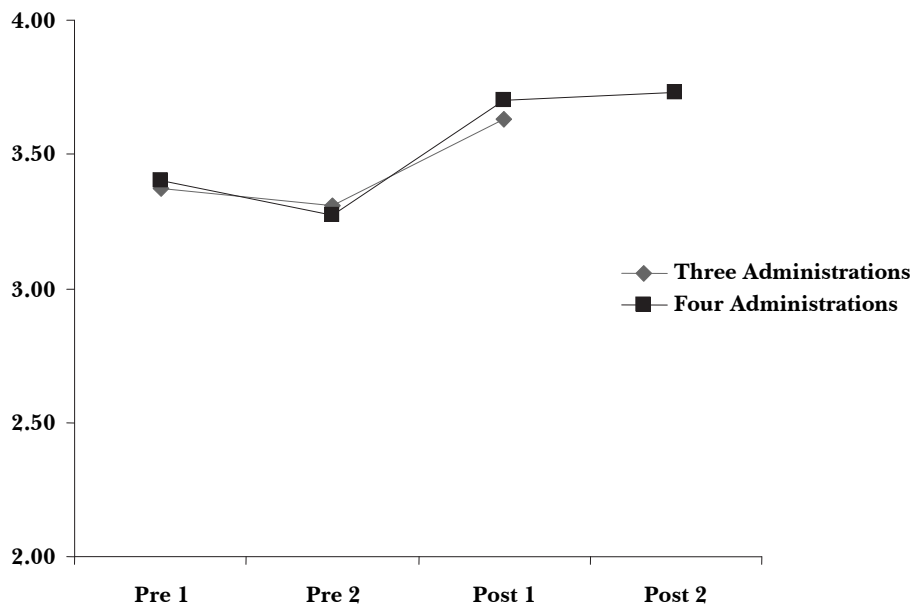


Table 5 shows the means and standard deviations of the ratings on the competencies selected for development in the second development cycle across multiple administrations of the multi-source feedback. The one-way ANOVA indicated a significant main effect of time for the group with three administrations ( $F(2, 108) = 25.07, p < 0.001$ ) and the group with four administrations ( $F(3, 45) = 15.14, p < 0.001$ ). The results of *post hoc* paired-sample t-tests show that for participants with three administrations, the Pre 2 mean rating on the selected competencies ( $M = 3.31$ ) was not statistically different

from the Pre 1 mean rating ( $M = 3.37$ ,  $p = 0.15$ ), but the Post 1 mean rating ( $M = 3.63$ ) was significantly higher than the Pre 2 mean rating ( $p < 0.001$ ). For participants with four administrations, the Pre 2 mean rating on the selected competencies was not statistically different from the Pre 1 mean rating ( $p > 0.05$ ). The Post 1 mean rating was significantly higher than the Pre 2 mean rating ( $M = 3.70$  and  $3.27$ , respectively;  $p < 0.001$ ). The Post 2 mean rating ( $M = 3.73$ ) was not statistically different from the Post 1 mean rating ( $M = 3.70$ ,  $p = 0.63$ ). Overall, the longitudinal analysis again indicated that the improvement after intervention was sustained but leveled off over time. Figure III depicts the mean ratings on the competencies selected for development in the second development cycle across multiple administrations of the multi-source feedback.

### Third Development Cycle

The means and standard deviations for the immediate pre- and post-assessment in the third development cycle are presented in Table 6. As can be observed, the mean rating on the competencies *not* selected for development and the overall composite score (i.e., the average of the 67 competencies) improved in the third development cycle, but were not statistically significant. In contrast, participants improved significantly on the competencies selected for development ( $p < 0.001$ ). The same method was used to test the regression to the mean. The observed difference between the mean rating on the competencies selected for development ( $M = 0.14$ ,  $s = 0.22$ ) was smaller than the expected difference ( $M = 0.19$ ,  $s = 0.11$ ). However, this difference was not statistically significant ( $p > 0.05$ ), because of the small sample size ( $N = 16$ ).

**Table 6**  
Immediate Pre- and Post-assessment in Third Development Cycle

	Pre		Post		<i>d</i>	<i>p</i> -value
	<i>Mean</i>	<i>Std.</i>	<i>Mean</i>	<i>Std.</i>		
Competencies Selected	3.49	0.32	3.77	0.26	0.93	$p < 0.001$
Competencies Not Selected	3.85	0.22	3.94	0.17	0.45	$p < 0.13$
All 67 Competencies	3.83	0.22	3.91	0.17	0.40	$p < 0.15$

*Note.*  $N = 16$ . *d* denotes effect size, the standardized mean difference.

**Table 7**  
Ratings on Competencies Selected in Third Development Cycle  
Across Multiple Administrations of the Feedback

Pre 1	Pre 2	Pre 3	Post 1
3.40 ( $s = 0.31$ )	3.45 ( $s = 0.30$ )	3.48 ( $s = 0.32$ )	3.77 ( $s = 0.26$ )

*Note.*  $N = 16$ .

These participants had four administrations of the multi-source feedback over a protracted period, including three pre-assessments and one post-assessment. Table 7 presents the means and standard deviations on the competencies selected for development in the third development cycle across the multiple administrations of the multi-source feedback. Again, the one-way ANOVA indicated a significant main effect of time ( $F(3, 45) = 7.69, p < 0.001$ ). On the other hand, the *post hoc* paired-sample t-tests showed that the three pre-assessments were not statistically different from each other ( $p > 0.05$ ). As predicted, the Post 1 assessment was statistically higher than the Pre 3 assessment ( $M = 3.77$  and  $3.48$ , respectively;  $p < 0.001$ ).

### Developmental Difficulty

Finally, this study tested whether the degree of developmental difficulty of the leadership competency itself may have influenced the feedback recipient's level of improvement. The average improvement on each of the competencies selected by the participants for development in the three development cycles was computed. This improvement was found to be negatively correlated with the leadership competency development difficulty index ( $r = -0.27, p < 0.05$ ) and development difficulty category ( $r = -0.31, p < 0.05$ ). Thus, as predicted, these findings suggest that participating managers made less improvement on the harder to develop leadership competencies than on the easier ones.

## DISCUSSION

Previous reviews of leadership development have indicated that the findings from empirical research examining the effectiveness of multi-source feedback were weak and inconsistent (Seifert *et al.*, 2003). This study sought to contribute to the literature by overcoming some of these methodological weaknesses of previous studies. First, in the current study, competencies specifically selected by managers were isolated and contrasted with changes in other competencies as well as changes in overall leadership. Second, changes were tracked over multiple administrations of the multi-source feedback. Therefore, this investigation truly represents a longitudinal examination of the impact of multi-source feedback over time. Finally, the "developmental difficulty" of various competencies was analyzed to determine if it influenced potential improvement.

Overall, the results consistently demonstrated that participants improved significantly *more* on competencies selected for development than on competencies *not* selected as well as than on a composite index of overall leadership. This improvement was greater than could be explained by the regression to the mean, and was sustained over the multiple administrations of the feedback. It also was found that the improvement in competencies was related to the developmental difficulty. As predicted, participants improved more on easier competencies than on harder ones. It presents the first study to investigate how the nature of the competency itself directly affects multi-source feedback effectiveness.

It is noteworthy that participants likewise improved on competencies *not* selected for development as well as a composite index of overall leadership. Although the gains often were statistically significant, the effect sizes were relatively small. For example, managers improved about 0.40 of one standard deviation on the composite leadership competency score. This effect size is higher than the ones observed by Smither *et al.*



(2005), but it is similar to the ones reported by Kluger and DeNisi (1996). On the other hand, participants improved more on those competencies selected for development. Across the three developmental cycles, the observed effect sizes were very consistent – between 0.93 and 0.97. According to Cohen (1988), these represent “large” effect sizes. Therefore, when the evaluation focuses on the direct objective of the developmental effort, one observes substantial improvement. This result is consistent with the findings from the research of other developmental programs. For example, De Meuse, Dai, and Lee (2009) found that when the effectiveness of executive coaching was evaluated, the literature indicated that individuals improved more in areas directly related to the coaching objective and less in other areas.

Researchers have proposed many theories as to why multi-source feedback may result in improvement on performance, such as Self-Consistency Theory (Korman, 1970, 1976; Johnson and Ferstl, 1999), Control Theory (Carver and Scheier, 1981), and Self-Awareness Theory (Atwater and Yammarino, 1992; Wegner and Vallacher, 1980). However, authors have debated over the years as to how to measure performance improvement. Many scholars have contrasted the post-feedback performance composite score with the pre-feedback composite score (as discussed previously). Some researchers have argued against the direct measure of the performance improvement (Buda *et al.*, 1991; Smither *et al.*, 1988). For example, Smither and Walker (2001) suggested a number of alternative approaches to measuring the impact of multi-source feedback. Built upon Smither *et al.*'s (2005) observation, the current study examined the amount of improvement on the leadership competencies participants selected for development following multi-source feedback. In all three development cycles, participants improved significantly more on the selected competencies than on competencies *not* selected for development as well as than on the composite competency score. The small effect sizes of multi-source feedback programs observed in previous research may be due (in part) to the method used to measure the performance improvement as opposed to the overall effectiveness of the multi-source feedback process.

Two additional findings of the current study are of interest. First, the improvement participants made on the competencies they selected for development was significantly greater than could be explained by the regression to the mean. Second, the improvement persisted across multiple administrations of the multi-source feedback. These findings are consistent with the results from Reilly *et al.* (1996). In their study, most of the performance improvement occurred between the first and second administration of the feedback program, and this improvement was not attributed solely to regression to the mean. Reilly *et al.* (1996) also observed that the improvement was sustained over a fairly lengthy period of time. The uniqueness of the current study was that it tested the regression to mean and the longitudinal effects of the multi-source feedback through multiple development cycles. In addition, it focused on the competencies participants selected for development rather than employing an overall index of leadership competency. The consistent findings in the current study clearly suggest that the multi-source feedback program achieved its objective to help managers develop their competencies.

Companies that use multi-source feedback will find it encouraging that the improvement had been sustained. On the other hand, this improvement happened largely in the early stages. Perhaps there is a ceiling effect. Managers may enhance their leadership competencies up to a certain level, and then it gets harder and harder to improve. It is reassuring to note that the regression to mean effect did not explain the amount of improvement observed in this study. Nevertheless, an alternative explanation

may be a “demand effect” leading to higher ratings after the feedback. For example, if an individual was asked to rate a manager at Time 1, and again at Time 2, it is possible that the individual may be biased. The individual, consciously or unconsciously, may rate the manager higher on the second administration, simply because he or she was asked to assess the improvement. The best way to rule out this alternative explanation is through the use of a control group design. Due to the fact that it was a field study, the employment of a control group was not possible in this instance. However, even without the use of a comparison group in this study, a demand effect is unlikely to explain the improvement observed. First, the multi-source feedback program implemented in this study had been operated under a high level of confidentiality. Other than the internal coach and the participating managers’ direct bosses, all other raters were blind to the competencies selected for development. Further, previous research has suggested an “assimilation effect” rather than the demand effect in feedback programs (Buda *et al.*, 1991; Smither *et al.*, 1988). An assimilation effect occurs when prior performance information biases rater expectations, which in turn, affect the rater’s cognitive categorization of the ratee. Thus, the recall of specific behaviors and performance needs to be consistent with the rater’s initial categorization. In the present study, the consistent results from multiple development cycles as well as the tests of the regression to mean support the overall effectiveness of the multi-source feedback.

Perhaps the most intriguing finding of the current study has to do with the relationship between the multi-source feedback effectiveness and leadership competency developmental difficulty. Studies investigating performance improvement within the context of multi-source feedback have not been conclusive (Bailey and Fletcher, 2002). The developmental outcomes following a multi-source feedback process likely depend on a complex interplay of factors (Craig and Hannum, 2006). In addition to the factors articulated by Smither *et al.* (2005), this study contends that the nature of the specific set of competencies selected for development is a critical factor influencing the developmental success following multi-source feedback. Bailey and Austin (2006) tested the indirect effects of the perceived importance of performance dimension. In their study, the perceived importance of feedback did not significantly moderate the association between feedback favorability and the amount of developmental activity or changes in self-assessments. The present study found that the developmental difficulty was statistically related to the observed improvement on the leadership competencies participants selected for development.

On the other hand, perhaps developmental difficulty is not the cause for little or no improvement in some competencies. Rather, improvement might be affected by how difficult it is to *observe* some competencies relative to other ones. Certainly, some competencies appear to be more concrete and descriptive, while others seem to be more abstract and less tangible (Viswesvaran *et al.*, 1996). For example, there is less chance for managers to observably demonstrate strategic thinking than interpersonal communication. In addition, strategic thinking represents a cognitive process that happens primarily in a manager’s mind; whereas, interpersonal communication has a high level of behavioral specificity. Future researchers should empirically test this distinction.

The findings of this study have several other theoretical and empirical implications. Theoretically, this study expands the knowledge base regarding situations in which feedback recipients are likely to make measurable performance improvement. Employees involved in a multi-source feedback program clearly improve most in the areas they work

on (i.e., leadership competencies identified in their IDPs). Nevertheless, many factors are likely to impact the success of such programs (Smither *et al.*, 2005). The results suggest that the nature of the competency itself is one of them. As one might suspect, employees appear to improve more on competencies which are easier to develop than on harder ones. Future models of feedback effectiveness should account for this effect. In addition, it is interesting (and heartening) to note that the recipients sustained this improvement over a number of years. Researchers recently have begun to explore how feedback and coaching can lead to long-term behavioral changes (Gray, 2006). Perhaps, when feedback recipients have a clear focus and work diligently, they are likely to make such sustained improvement. This finding needs to be replicated in future longitudinal research before there can be confidence in the long-term efficacy of multi-source feedback.

From a practical perspective, it can help coaches and feedback recipients decide on the number of developmental goals to include in the IDP. A general rule of thumb is that feedback recipients should focus on no more than two or three areas for improvement (Antonioni, 1996). The findings of the current study imply that this number probably depends on the developmental difficulty of the competencies selected. Feedback recipients may decide to work on a greater or fewer number of competencies depending on how difficult the competencies selected are to develop. Obviously, feedback recipients also may want to consider other factors to determine the number of competencies for development. Even for easy to develop competencies, managers may find it problematic if they select too many objectives. Another practical implication is that organizations should consider providing additional resources to support feedback recipients when harder competencies are chosen. The awareness of the inherent nature of the leadership competency can help organizations design action plans and provide supporting resources and coaching that maximize multi-source feedback effectiveness.

Finally, program effectiveness should not be confused with program impact when evaluating multi-source feedback. A manager may make only a small improvement on difficult competencies (e.g., strategic and visioning skills). These slight improvements, however, may have a huge impact on organizational performance. In contrast, a substantial change on some relatively easy to develop competencies might have very little impact on overall business results. This study found a relationship between the effectiveness of the multi-source feedback and the developmental difficulty of the leadership competency. Readers should not generalize these findings to the overall impact or ROI (return on investment) of such a feedback program.

An obvious limitation of the study is the relatively small sample size. In addition, the overall sample size decreased across the three development cycles. However, the analyses suggested that the attrition had no impact on the results. Further, the longitudinal results from the three development cycles were very consistent. Another limitation of this study is that the effects of the multi-source feedback might have been confounded by other factors. As described in the method section, the company provided coaching as well as implemented other interventions to hold the participating managers accountable. Multi-source feedback (combined with the coaching) may have led to the improved effectiveness (Luthans and Peterson, 2003; Smither *et al.*, 2003). Consequently, it is difficult to discern whether the positive effects observed in this study were due to the multi-source feedback or other organizational practices. Nevertheless, the findings in the current study are encouraging to companies using multi-source feedback. Organizations rarely use multi-source feedback in isolation. At a minimum, a feedback facilitator should be provided to help recipients interpret the feedback data (Seifert *et al.*, 2003). Other organizational

practices, particularly executive coaching, frequently are implemented together with multi-source feedback. Without the support of these proactive organizational efforts and an organizational culture focused on leadership development, multi-source feedback programs are unlikely to be successful.

The findings from the present study largely support the use of multi-source feedback. Human resource professionals and other administrators of multi-source feedback programs can be reassured that multi-source feedback programs do appear to work. The data clearly demonstrate that managers show improvement on their competencies following such feedback. Further, the results demonstrate that the gains are sustained for several years. However, the findings also suggest that the efficacy of these feedback programs is affected by (1) the developmental difficulty of the competencies the managers select to develop and (2) how an organization measures competency improvement. It would seem prudent for companies to carefully consider these factors when evaluating the success of such programs. Likewise, scholars should recognize that previous research may systematically underestimate the effectiveness of multi-source feedback programs. Future field research is needed to support and refine these findings.

### References

- Albright, M. D. and P. E. Levy. 1995. "The Effects of Source and Performance Rating Discrepancy on Reactions to Multiple Raters." *Journal of Applied Social Psychology* 25: 577-600.
- Antonioni, D. 1996. "Designing an Effective 360-Degree Appraisal Feedback Process." *Organizational Dynamics* 25 (2): 24-38.
- Atwater, L. E. and J. F. Brett. 2005. "Antecedents and Consequences of Reactions to Developmental 360-Degree Feedback." *Journal of Vocational Behavior* 66: 532-548.
- \_\_\_\_\_ and F. Yammarino. 1992. "Does Self-Other Agreement on Leadership Perceptions Moderate the Validity of Leadership and Performance Predictions?" *Personnel Psychology* 45: 141-164.
- Bailey, C. and M. Austin. 2006. "360-Degree Feedback and Developmental Outcomes: The Role of Feedback Characteristics, Self-efficacy and Importance of Feedback Dimensions to Focal Managers' Current Role." *International Journal of Selection and Assessment* 14: 51-66.
- \_\_\_\_\_ and C. Fletcher. 2002. "The Impact of Multiple Source Feedback on Management Development: Findings from a Longitudinal Study." *Journal of Organizational Behavior* 23: 853-867.
- Becker, B. E., M. A. Huselid, and D. Ulrich. 2001. *The HR Scorecard: Linking People, Strategy, and Performance*. Boston, MA: Harvard Business School Press.
- Borman, W. C. 1997. "360-Degree Ratings: An Analysis of Assumptions and a Research Agenda for Evaluating Their Validity." *Human Resource Management Review* 7: 299-315.
- Brutus, S., M. Derayeh, C. Fletcher, C. Bailey, P. Velazquez, K. Shi, C. Simon, and V. Labath. 2006. "Internationalization of Multi-source Feedback Systems: A Six-country Exploratory Analysis of 360-degree Feedback." *International Journal of Human Resource Management* 17: 1888-1906.
- Buda, R., R. R. Reilly, and J. W. Smither. 1991. "The Influence of Indirect Knowledge of Prior Performance on Evaluations of Present Performance: The Generalizability of

- Assimilation Effects." *Journal of Psychology and the Behavioral Science* 6: 89-99.
- Carruthers, F. 2003. "Nothing but the Truth." *Australian Financial Review* (November 14): 78.
- Carver, C. and M. Scheier. 1981. *Attention and Self-Regulation: A Control-Theory Approach to Human Behavior*. New York, NY: Springer.
- Cohen, J. 1988. *Statistic Power Analysis for the Behavioral Science* (2<sup>nd</sup> ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
- Craig, S. B. and K. Hannum. 2006. "Research Update: 360-Degree Performance Assessment." *Consulting Psychology Journal: Practice and Research* 58: 117-124.
- De Meuse, K. P., G. Dai, and R. J. Lee. 2009. "Evaluating the Effectiveness of Executive Coaching: Beyond ROI?" *Coaching: An International Journal of Theory, Research and Practice* 2 (2): 118-135.
- Dunning, D., C. Heath, and J. M. Suls. 2004. "Flawed Self-assessment: Implications for Health, Education, and the Workplace." *Psychological Science in the Public Interest* 5: 69-106.
- Edwards, M. R. and A. J. Ewen. 1996. *360-Degree Feedback: The Powerful New Model for Employee Assessment and Performance Improvement*. New York, NY: American Management Association.
- Eichinger, R. W. and M. M. Lombardo. 2004. "Patterns of Rater Accuracy in 360-Degree Feedback." *Human Resource Planning* 27: 23-25.
- Fletcher, C. and C. Baldry. 2000. "A Study of Individual Differences and Self-awareness in the Context of Multi-Source Feedback." *Journal of Occupational and Organizational Psychology* 73: 303-319.
- \_\_\_\_\_. and \_\_\_\_\_. 1999. "Multi-source Feedback: A Research Perspective." In *International Review of Industrial and Organizational Psychology, Vol. 14*. Eds. C. L. Cooper and I. T. Robertson. London: John Wiley.
- Gray, D. E. 2006. "Executive Coaching: Towards a Dynamic Alliance of Psychotherapy and Transformative Learning Process." *Management Learning* 37: 475-497.
- Harris, M. M. and J. Schaubroeck. 1988. "A Meta-Analysis of Self-Supervisor, Self-Peer, and Peer-Supervisor Ratings." *Personnel Psychology* 41: 43-62.
- Hegarty, W. H. 1974. "Using Subordinate Ratings to Elicit Behavior Changes in Supervisors." *Journal of Applied Psychology* 59: 764-766.
- Heslin, P. A. and G. P. Latham. 2004. "The Effect of Upward Feedback on Managerial Behavior." *Applied Psychology: An International Review* 53: 23-37.
- Johnson, J. W. and K. L. Ferstl. 1999. "The Effects of Interrater and Self-Other Agreement on Performance Improvement Following Upward Feedback." *Personnel Psychology* 52: 271-303.
- Kluger, A. N. and A. DeNisi. 1996. "The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory." *Psychological Bulletin* 119: 254-284.
- Korman, A. K. 1976. "Hypothesis of Work Behavior Revisited and an Extension." *Academy of Management Review* 1: 50-63.
- \_\_\_\_\_. 1970. "Toward a Hypothesis of Work Behavior." *Journal of Applied Psychology* 54: 31-41.
- Lawler, E. E. 1967. "The Multitrait-Multirater Approach to Measuring Managerial Job Performance." *Journal of Applied Psychology* 51: 369-380.
- Lievens, F., J. I. Sanchez, and W. de Corte. 2004. "Easing the Inferential Leap in Competency Modeling: The Effects of Task-related Information and Subject Matter

- Expertise." *Personnel Psychology* 57: 881-904.
- Lombardo, M. M. and R. W. Eichinger. 2004. *The Leadership Machine*. Minneapolis, MN: Lominger Limited.
- \_\_\_\_\_, and \_\_\_\_\_. 2003. *Leadership Architect Norms and Validity Report*. Minneapolis, MN: Lominger Limited.
- \_\_\_\_\_, and \_\_\_\_\_. 1996. *FYI For Your Improvement: A Guide for Development and Coaching*. Minneapolis, MN: Lominger Limited.
- \_\_\_\_\_, and \_\_\_\_\_. 1995. *The CAREER ARCHITECT® User's Manual*. Minneapolis, MN: Lominger Limited.
- \_\_\_\_\_, and \_\_\_\_\_. 1989. *Eighty-eight Assignments for Development in Place*. Greensboro, NC: Center for Creative Leadership.
- London, M. and J. W. Smither. 1995. "Can Multi-source Feedback Change Perceptions of Goal Accomplishment, Self-evaluations, and Performance-related Outcomes? Theory Based Applications and Directions for Research." *Personnel Psychology* 48: 803-839.
- Luthans, F. and S. J. Peterson. 2003. "360-Degree Feedback with Systematic Coaching: Empirical Analysis Suggests a Winning Combination." *Human Resource Management* 42: 243-256.
- Morgeson, F. P., T. V. Mumford, and M. A. Campion. 2005. "Coming Full Circle: Using Research to Address 27 Questions about 360-Degree Feedback Programs." *Consulting Psychology Journal: Practice and Research* 57: 196-209.
- Mount, M. K., T. A. Judge, S. E. Scullen, M. R. Sytsma, and S. A. Hezlett. 1998. "Trait, Rater, and Level Effects in 360-Degree Performance Ratings." *Personnel Psychology* 51: 557-576.
- Pfau, B. and I. Kay. 2002. "Does 360-degree Feedback Negatively Affect Company Performance?" *HR Magazine*. Retrieved from [http://findarticles.com/p/articles/mi\\_m3495/is\\_6\\_47/ai\\_87461017/](http://findarticles.com/p/articles/mi_m3495/is_6_47/ai_87461017/) (accessed June 5, 2009).
- Reilly, R. R., J. W. Smither, and N. L. Vasilopoulos. 1996. "A Longitudinal Study of Upward Feedback." *Personnel Psychology* 49: 599-612.
- Schippmann, J. S., R. A. Ash, M. Battista, L. Carr, L. D. Eyde, B. Hesketh, J. Kehoe, K. Pearlman, E. P. Prien, and J. I. Sanchez. 2000. "The Practice of Competency Modeling." *Personnel Psychology* 53: 703-740.
- Scullen, S. E., M. K. Mount, and M. Goff. 2000. "Understanding the Latent Structure of Job Performance Ratings." *Journal of Applied Psychology* 85: 956-970.
- Seifert, C. F., G. Yukl, and R. A. McDonald. 2003. "Effects of Multi-source Feedback and a Feedback Facilitator on the Influence Behavior of Managers toward Subordinates." *Journal of Applied Psychology* 88: 561-569.
- Smither, J. W., M. London, R. Flautt, Y. Vargas, and I. Kucine. 2003. "Can Working with an Executive Coach Improve Multi-source Feedback Ratings Over Time? A Quasi-Experimental Field Study." *Personnel Psychology* 56: 23-44.
- \_\_\_\_\_, \_\_\_\_\_, and R. R. Reilly. 2005. "Does Performance Improve Following Multisource Feedback? A Theoretical Model, Meta-analysis, and a Review of Empirical Findings." *Personnel Psychology* 58: 33-66.
- \_\_\_\_\_, R. R. Reilly, and R. Buda. 1988. "The Effect of Prior Performance Information on Rating of Present Performance: Contrast Versus Assimilation Revisited." *Journal of Applied Psychology* 73: 487-496.
- \_\_\_\_\_, and A. G. Walker. 2001. "Measuring the Impact of Multisource Feedback." In *Handbook of Multisource Feedback*. Eds. D. Bracken, C. Timmreck, and A. Church. San

- Francisco, CA: Jossey-Bass.
- Spencer, L., D. McClelland, and S. Spencer. 1994. *Competency Assessment Methods: History and State of the Art*. Boston, MA: Halt/McBer Research Press.
- Tett, R. P., H. A. Guterman, A. Bleier, and P. J. Murphy. 2000. "Development and Content Validation of a 'Hyperdimensional' Taxonomy of Managerial Competence." *Human Performance* 13: 205-251.
- Viswesvaran, C., D. S. Ones, and F. L. Schmidt. 1996. "Comparative Analysis of the Reliability of Job Performance Ratings." *Journal of Applied Psychology* 81: 557-574.
- Walker, A. G. and J. W. Smither. 1999. "A Five-Year Study of Upward Feedback: What Managers Do with Their Results Matters." *Personnel Psychology* 52: 393-423.
- Wegner, D. and R. Vallacher. 1980. *The Self in Social Psychology*. New York, NY: Oxford University Press.
- Wohlers, R. J. and M. London. 1989. "Ratings of Managerial Characteristics: Evaluation Difficulty, Co-worker Agreement, and Self-awareness." *Personnel Psychology* 42: 235-261.

Copyright of Journal of Managerial Issues is the property of Journal of Managerial Issues / PSU and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.